Coupling ensemble Kalman filter with four-dimensional variational data assimilation

Fuqing Zhang and Meng Zhang

Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

James A. Hansen

Navy Research Laboratory, Monterrey, California

Submitted to *Monthly Weather Review* for Publications

22 April 2007

Corresponding author address: Dr. Fuqing Zhang, Department of Atmospheric Sciences, Texas A&M University, College Station, TX 77845-3150.

E-mail: fzhang@tamu.edu

Abstract

This study examines the performance of coupling deterministic four-dimensional variational assimilation (4DVAR) with an ensemble Kalman filter (EnKF) to produce a superior hybrid approach for data assimilation. The coupled assimilation scheme (E4DVAR) benefits from using the state-dependent uncertainty provided by EnKF while taking advantage of dynamic constraints in 4DVAR that aid in preventing filter divergence. The deterministic analysis produced by 4DVAR provides an estimate of the minimum error variance state about which the ensemble perturbations are transformed, and the resulting ensemble analysis can be propagated forward both for the next assimilation cycle and as a basis for ensemble forecasting. The feasibility and effectiveness of this coupled approach are demonstrated in an idealized model with simulated observations. It is found that the E4DVAR is capable of outperforming both 4DVAR and the EnKF under both perfect- and imperfect-model scenarios. The performance of the coupled scheme is also less sensitive to either the ensemble size or the assimilation window length than that for standard EnKF or 4DVAR implementations.

1 Introduction

Data assimilation is the blending of two independent estimates of the state of a system, typically in the form of observational information and a short-term model forecast, in a manner consistent with their respective uncertainties (Talagrand 1997). Data assimilation can be broadly classified as deterministic versus ensemble-based, and as local in time vs. distributed in time. Ensemble Kalman filters (EnKF) are ensemble-based, local in time approaches to data assimilation (Evensen 1994) while four-dimensional variational assimilation (4DVAR) is deterministic, distributed in time (Courtier et al. 1994). Recently, Caya et al. (2005) directly compared these two current state-of-the-art data assimilation techniques for storm-scale data assimilation, and clearly demonstrated the strengths and weaknesses of each approach. In a perfect-model setting, they found that the dynamic constraints imposed by 4DVAR were able to generate good, dynamically consistent analyses almost immediately. It took longer for the EnKF to spin up, but ultimately the state-dependent uncertainty information utilized by the EnKF enabled it to outperform 4DVAR (in terms of root-mean square error or RMSE), which used very simplistic first guess information. The current study seeks to advance the state-of-the-science in data assimilation by coupling 4DVAR with EnKF aiming at maximally exploiting the strengths of the two forms of data assimilation, while simultaneously offsetting their respective weaknesses. A hybrid form of the ensemble-based methods using three-dimensional variational data assimilation (3DVAR) has been previously used in Hamill and Snyder (2000) and more recently Wang et al. (2007). To a broader extent, the Houtekamer et al. (2001) concept of additively combining ensemble-based covariance estimates with those from a 3DVAR background error covariance can be regarded as a special form of a hybrid approach. The current work can be viewed as an extension to previously published hybrid methods.

2 EnKF, 4DVAR and E4DVAR

2.1 EnKF

Ensemble (Monte-Carlo) approximations to the extended Kalman filter use sample statistics to define the uncertainty information associated with the prior state estimate. Define $\overline{\mathbf{x}}^f \in \Re^n$ to be the prior minimum error variance estimate of the state, and $\mathbf{P}^f$ to be the covariance matrix that defines the uncertainty associated with the prior. An estimate of $\mathbf{P}^f$ is obtained by considering k ensemble members, $\mathbf{x}_i^f, i = 1, k$, such that $\overline{\mathbf{x}}^f = \frac{1}{k}\sum_{i=1}^{k}\mathbf{x}_i^f$, and

$$\mathbf{P}^f \cong \frac{1}{k-1}\left(\mathbf{X}^f - \overline{\mathbf{X}}^f\right)\left(\mathbf{X}^f - \overline{\mathbf{X}}^f\right)^T, \qquad (1)$$

where $\mathbf{X}^f$ is an n by k matrix where each column is an ensemble member, $\mathbf{x}_i^f$, and $\overline{\mathbf{X}}^f$ is an n by k matrix where each column is the ensemble mean, $\overline{\mathbf{x}}^f$. Given this prior information, and assuming observations, $\mathbf{y}$, and their error covariance, $\mathbf{R}$, are available, the posterior minimum error variance estimate of the state (the analysis) is $\overline{\mathbf{x}}^a = \frac{1}{k}\sum_{i=1}^{k}\mathbf{x}_i^a$ in that $\mathbf{x}_i^a$ is given by

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_i^f), \qquad (4)$$

where $\mathbf{H}$ is is an observation operator that maps from model space to observation space. Also, the expected posterior uncertainty is given by

$$\mathbf{P}^a = \mathbf{P}^f - \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{H}\mathbf{P}^f. \qquad (5)$$

There are several variations to the original ensemble Kalman filter (EnKF) first proposed Evensen (1994) and later in Houtekamer and Mitchell (1998) including the use of the ensemble square root filter (EnSRF, Whitaker and Hamill 2002; Snyder and Zhang 2003), the ensemble adjustment filter (EAF, Anderson, 2001), and the ensemble transform Kalman filter (ETKF, Bishop et al 2001). In this work the EnSRF-version of the EnKF is used.

2.2 4DVAR

Data assimilation via 4DVAR proceeds through the minimization of a cost function containing observations that are distributed in time and a background estimate. The traditional 4DVAR cost function can be written as

$$J(\mathbf{x}_0) = (\mathbf{x}^b - \mathbf{x}_0)^T \mathbf{B}^{-1} (\mathbf{x}^b - \mathbf{x}_0) + \sum_{n=0}^{N} (\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_n)^T \mathbf{R}_n^{-1} (\mathbf{y}_n - \mathbf{H}_n \mathbf{x}_n) \qquad (6)$$

where $\mathbf{x}^b$ is the first guess at the system state (the equivalent of $\overline{\mathbf{x}}^f$ in the ensemble filter discussion above) , $\mathbf{B}$ is the background error covariance defining the uncertainty associated with the first guess (the equivalent of $\mathbf{P}^f$ in the ensemble filter discussion above), $\mathbf{y}_n$ is an observation at time *n*, $\mathbf{H}_n$ and $\mathbf{R}_n$ are the associated observation operator and error covariance, and the $\mathbf{x}_n$ are the model estimates of the system state through the assimilation window. Data assimilation proceeds by adjusting the initial condition $\mathbf{x}_0$ to $\mathbf{x}_0^{optimal}$, so that when $\mathbf{x}_n^{optimal}$ propagates forward in time it gets as close as possible to the observations $\mathbf{y}_n$ in assimilation window *N*, conditional upon $\mathbf{x}_0^{optimal}$ not getting too far from the first guess value, $\mathbf{x}^b$. Here "close" and "too far" are defined by the background and observation covariance matrices, $\mathbf{B}$ and $\mathbf{R}_n$.

As with the ensemble-based filters, there are numerous approaches to estimating the minimum of the cost function in equation 6. In this work we employ a limited-memory quasi-Newton method (L-BFGS) (Liu and Nocedal 1989) for the minimization in all 4DVAR approaches. The L-BFGS method is found to have superb performance in nonlinear minimization problems and has relatively low computing cost and low storage requirement.

2.3 E4DVAR: coupling the EnKF and 4DVAR

Conceptually, the coupled approach, hereafter termed as "E4DVAR", aims to link the distributed in time deterministic approach of 4DVAR and the local in time ensemble-based approach of the EnKF. However, while the ensemble-based filters benefit from their use of state-dependent uncertainty information and from the explicit and consistent production of analysis ensembles for forecasting, limited ensemble sizes render the sample covariances rank deficient and inaccurate, which would result in bad ensemble analyses due to filter divergence. Rather ad hoc fixes such as localization (Gaspari and Cohn, 1999) are applied to the covariance in order to increase the rank of the prior covariance. Deterministic 4DVAR, on the other hand, benefits from the dynamic constraint of finding a model trajectory that gets as close as possible to a trajectory of observations, enabling it to overcome rather uninformative, static background error covariance information. In addition to the limitations of the static background error covariance, another limitation of deterministic 4DVAR in the forecast system sense is the lack of the production of an internally consistent analysis ensemble. The proposed E4DVAR data assimilation scheme uses the respective strengths of the two constituent schemes to off-set the weaknesses of each: the state-dependent uncertainty information and ensemble construction capability of the ensemble-based filter compensates for the inherent weaknesses of 4DVAR, while the ability of 4DVAR to overcome inaccuracies in the background error covariance to produce a good approximation to the minimum error covariance estimate of the system state compensates for an inherent weakness of the ensemble-based filter.

There are many possible implementations of E4DVAR but for the purpose of clarity we choose to concentrate on a representative formulation. The mechanics of this representative scheme couples 4DVAR with an EnKF where the state and perturbation updates have been separated. An illustration of the E4DVAR coupling procedure used in the current study is depicted in the schematic flowchart of Figure 1: a prior ensemble forecast produced by the EnKF that is valid at time t is used to estimate $\mathbf{P}^f$ for the subsequent 4DVAR assimilation cycle (t=j,j+1) while the 4DVAR analysis from the previous assimilation cycle (t=j-1,j) is used to replace the EnKF analysis mean for subsequent ensemble forecast. More generally, if there are observations between t=(j, j+1), the standard EnKF will be used to assimilate those observations (that will be within the dotted box of labeled with "Ensemble forecast" in Fig. 1). An alternative stronger coupling is to replace the posterior ensemble mean with the 4DVAR trajectory after each EnKF analysis. Also, for the benefits to be clear in section 4, we can further mix the static and the ensemble-estimated flow-dependent background error covariance (Hamill and Snyder 2000) through

$$\mathbf{B} = \beta \mathbf{P}^f + (1-\beta)\mathbf{B}_s \tag{7}$$

where the mixing coefficient β is the weight given to the ensemble covariance estimate, in which the E4DVARis same as the standard 4DVAR for β=0 while all background error covariance comes from ensembles for β=1 and in between for β=0.5.

3 Experimental design

This proof-of-concept study will be carried out using the model of Lorenz (1996):

$$\frac{dx_i}{dt} = -x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + F, \quad i = 1, n, \tag{8}$$

with cyclic boundary conditions. Although not derived from any known fluids equations, the dynamics of equation (8) are "atmosphere-like" in that they consist of nonlinear advection-like terms, a damping term, and an external forcing; they can be thought of as some atmospheric quantity distributed on a latitude circle. One can choose any dimension, *n*, greater than 4 and obtain chaotic behavior for suitable values of *F*. The base-line configuration was $n = 80$ and $F = 8$, which is computationally stable with a time step of 0.05 units, or 6 h in equivalent.

The performance of two coupled approaches of E4DVar, one with β=1 and the other with β=0.5 (see eq. 7), is examined in comparison to the standard non-coupled methods (EnKF and 4DVAR). Ensemble sizes ranging between $k = 10$ and $k = 500$ were considered in the experiments utilizing ensemble techniques but most results were shown for $k = 40$ and $k = 10$. The default number of observations is $m = 20$ (equally spaced). Observations were taken every 2 steps, or 12 h (as for standard soundings), and specified observational error of 0.2 that is approximately 3% the radius of the attractor. For 4DVAR, we considered the assimilation window length of both $N = 4$ (standard 24-h daily assimilation cycle) and $N = 20$ (near optimum window of 60h for this dynamic system studied). The standard 4DVAR uses a diagonal background error covariance whose values (all equal to 0.04) were determined through the statistics of a long EnKF integration.

The covariance relaxation method of Zhang et al. (2004)

$$(\mathbf{x}_i^{'})^{new} = \alpha(\mathbf{x}_i^{'})^f + (1-\alpha)(\mathbf{x}_i^{'})^a \qquad (9)$$

and the covariance localization based on Gaspari and Cohn (1999) will be used for all ensemble-based experiments. Other methods of boosting and covariance localization radius were also assessed but did not yield better performance (not shown). All experiments were carried out over 10 years, and assessment took place through comparison of ensemble mean analysis errors in the full model space. We recognize that the specific details of our implementations are unlikely to have direct relevance to models of real systems, so we pay particular attention to sensitivity experiments that help us identify the conditions under which various assimilation schemes perform well.

4 Results

4.1 Perfect-model experiments

Figure 2 compares the performance of the coupled approach (two E4DVAR implementations with β=0 and 0.5, respectively) with the standard EnKF and 4DVAR under the perfect-model assumption (F=8 for all truth, forecast and assimilation runs) and for the assimilation window length 20 and an ensemble size of 40 and 10, respectively. A radius of influence of 8 and a relaxation coefficient of $\alpha = 0.5$ are used for all ensemble experiments. It is clear from Figure 1 that, without model error and given typical ensemble size ($k = 40$), all methods will give satisfactory performance in terms of overall RMS error, in which all methods with ensemble-based flow-dependent background error covariances are slightly better than standard 4DVAR with static B (Fig. 2a). Remarkably, with a reduced ensemble size of $k = 10$, degradation in the performance of the coupled approaches is rather insignificant while the standard EnKF fails quickly because of filter divergence (Fig. 2b).

However, an acceptable performance of the standard EnKF with $k = 10$ may still be achieved with a smaller radius of influence (R=4) and relaxation the error covariance more to the prior $\alpha = 0.7$. Some small improvement can also be achieved for other ensemble-based experiments through using different localization radius, relaxation and mixing coefficients (Table 1 & 2). Noticeably, when a large ensemble size is used, the ensemble methods will benefit more from using a larger radius of influence, smaller relaxation coefficient and a larger mixing coefficient, which is consistent with a smaller sampling error in the ensemble-based covariance estimate. Tuning the static B through varying the covariance magnitude does not yield improvement for the standard 4DVAR but it is very sensitive to the assimilation window length. Significant degradation in 4DVAR performance is observed if a standard 24-h (shorter) assimilation window id used (Table 2), partly due to frequent encountering of local minima in its minimization (not shown), much more than those in Fig. 2 (e.g., a RMSE spike during year 3-4). Also, the advantage of the coupled approach may be more (less) pronounced if less (more) observations are assimilated (not shown).

4.2 Experiments with moderate model error

In these experiments, the forecast model in all assimilation methods used a different (incorrectly-specified) forcing coefficient (F=8.5) from that used in the truth simulation (F=8.0). The truth run is used for verification and for generating observations. The ensemble-mean derivation (with model error, F=8.5) from the perfect-model ensembles (F=8.0) over 24 h (starting from the same initial perturbations every 24h and averaged over 10 years) is approximately 20% and 30% of the forecast ensemble spread of 40 and 10 members, respectively.

Figure 3 shows the performance of the EnKF, 4DVAR and the coupled approaches with an imperfect forecast model (F=8.5) for different ensemble sizes. The experiment configurations are exactly the same as those for the perfect model (Fig. 2) except that a radius of influence of 4 (vs. 8) and a relaxation coefficient of $\alpha = 0.6$ (vs. 0.5) are used for all associated experiments. The use of a smaller radius of influence and larger relaxation coefficient are a direct consequence of degradation of the ensemble-based error covariance estimate in the presence of model error. With moderate model error and an ensemble size of $k = 40$, all methods will still give satisfactory performance (values below 1.0 or 20-25% of the climatological uncertainty), though each of them will have significantly larger overall RMS error than the corresponding perfect-model experiments (Fig. 3 vs. Fig. 2; Table 1 & 2). Noticeably, in the presence of moderate model error, the standard 4DVAR performs significantly better than EnKF for an assimilation window of 60 h (N=10) (Fig. 3, Table 1) and the advantage of using the standard EnKF over the standard 4DVAR become much smaller for an assimilation window of 24 h (N=4) (Table 2), both of which are inferior to the two coupled approaches. Even with an ensemble size of 10, both coupled approaches can perform considerably better than 4DVAR, but in this case, significantly better performance is achieved through mixing the flow-dependent and static error covariance, which reduced both the appropriate and inappropriate correlations and prevented the underestimation of background error variance (Table 1 & 2).

With an ensemble size of 10, the EnKF may barely function without filter divergence (though performs poorly) with an even smaller radius of influence (R=3) and a stronger relaxation of the error covariance to the prior with a mixing coefficient of $\alpha = 0.7$ (Table 1). Again, some small improvement can also be achieved for other ensemble-based experiments through using different localization radius, relaxation and mixing coefficients (Table 2). These additional sensitivity experiments demonstrate that, when an imperfect model is used, the ensemble methods will benefit more from using a smaller radius of influence, a larger relaxation coefficient and a smaller mixing coefficient, which is consistent with the degradation of the quality of the ensemble-based error covariance estimate (Hansen 2002; Meng and Zhang 2007).

4.3 Experiments with severe model error

In these experiments, the forecast model in all assimilation methods used a different (incorrectly-specified) forcing coefficient (F=9.0) from that used in the truth simulation (F=8.0). The ensemble-mean derivation (with model error, F=9.0) from the perfect-model ensembles (F=8.0) over 24 h (starting from the same initial perturbations every 24h and averaged over 10 years) is approximately 35% and 50% of the forecast ensemble spread of 40 and 10 members, respectively.

Figure 4 shows the performance of data assimilation methods will suffer greatly if the forecast model is fundamentally flawed. In this case, the standard 4DVAR will have an unacceptable overall RMSE of 1.12 for an assimilation window of 60 h or N=10 and an unacceptable overall RMS error of 1.52 for a shorter assimilation window of 24 h while the standard EnKF with radius of influence (R=4) will not converge at all. However, an acceptable performance can still be achieved with the coupled approaches, especially through mixing the flow-dependent and static error covariance, even with an ensemble size of 10 (Fig. 4). With such severe model error, stronger sensitivity is found for the ensemble methods and thus more delicate tuning is necessary (Table 1 & 2).

Results from these imperfect-model experiments imply that, while model error imposes strong limitation on all data assimilation approached, the use of stronger dynamic constraints in 4DVAR makes it less vulnerable to model errors than EnKF.

5 Concluding remarks

We have found the coupled data assimilation approach (E4DVAR) to be effective in the context of an idealized model; the coupled approach is able to produce analyses that are superior to those produced either by the standard EnKF or 4DVAR under both perfect imperfect model scenarios. Extensive sensitivity studies using the idealized model have helped to elucidate when and why the coupled approaches are effective. In this context, 4DVAR's primary strength is the model dynamic constraint that has an ability to overcome relatively uninformative information about the uncertainty associated with the first guess of the system state, but its primary weaknesses are the poor initial uncertainty estimates and the production of a single, deterministic state estimate. The primary strengths of the EnKF are a state-dependent estimate of first guess uncertainty (provided by an ensemble forecast) and the production of an ensemble of analyzed initial conditions. Its primary weakness is an extreme sensitivity to the quality of the state-dependent estimate of uncertainty. The coupled schemes use the respective strengths of the two constituent schemes to off-set the weaknesses of each: the state-dependent uncertainty information and ensemble construction capability of the ensemble filter addresses the inherent weaknesses of 4DVAR, while the ability of 4DVAR to overcome inaccuracies in the background error covariance to produce a good approximation to the minimum error variance estimate of the system state addresses an inherent weakness of the ensemble-based filter.

One should never expect individual results from simplified models to have any relevance for more complex models. However, one should also not expect that issues elucidated in the context of a simplified model to simply disappear when more complex models are considered. We therefore anticipate that the proposed coupled approach to data assimilation will be fruitful for models of "real" systems in some regions of parameter space (assimilation window length, observation distribution, observation frequency, observation error level, ensemble size); we can not know *a priori* if those regions will correspond to the area of parameter space defined by current operational constraints and we can not know *a priori* if the improvement will balance the increase in computational cost. For the current study, the computational cost of the coupled approach is slightly higher than the sum of the standard EnKF and 4Dvar, partly due to the trivial inexpensive inversion of a simple diagonal B matrix for the standard 4Dvar. We envision in real-data atmospheric applications, the difference of computational cost between E4DVAR and the two standard approaches (4DAVR and EnKF) will be much reduced since the coupled approach allows the use of a smaller ensemble size while the use of follow-dependent B may reduce the number of minimization iterations.

Acknowledgments: This research is sponsored by the NSF grant ATM0205599 and by the Office of Navy Research under grant N000140410471.

References

Anderson, J.L, 2001: An Ensemble Adjustment Kalman Filter for Data Assimilation. *Mon. Wea. Rev.*, **129**, 2284-2903.

Bishop, C.H., B.J. Etherton, and S.J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.

Caya A., J. Sun, and C. Snyder, 2005: A comparison between the 4Dvar and the ensemble Kalman filter for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081-3094.

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120,** 1367–1387.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143-10162.

Gaspari, G., and S. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Q. J. R. Meteor. Soc.*, **125**, 723-757.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

Hansen, J. A., 2002: Accounting for model error in ensemble-based state estimation and forecasting. *Mon. Wea. Rev.*, **130**, 2373-2391.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.

Houtekamer P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev*, **129**, 123–137.

Hansen, J. A., 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604-620.

Liu, D. C., and Nocedal, J., 1989, On the limited memory BFGS method for large scale optimization: *Mathematical Programming*, **45**, 503-528.

Lorenz, E. N., 1996: Predictability – A problem partly solved. In “Predictability”. ECMWF, Seminar Proceedings, Shinfield Park, Reading, RG2 9AX, 1-18.

Meng, Z., and F. Zhang, 2007: Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*,, **135**, 1403-1423.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663-1677.

Talagrand, O., 1997: Assimilation of observations, an introduction. *J. Meteor. Soc. Japan*, **75,** 191–209.

Wang X., C. Snyder, T. M. Hamill, 2007: On the Theoretical Equivalence of Differently Proposed Ensemble–3DVAR Hybrid Analysis Schemes. Mon. Wea. Rev.: 135, 222–227

Whitaker, J. S., and T.M. Hamill, 2002: Ensemble Data Assimilation Without Perturbed Observations. *Mon. Wea. Rev.*, **130**, 1923.

Zhang, F., C. Snyder, and J. Sun, 2004: Tests of an ensemble Kalman filter for convective-scale data assimilation: Impact of initial estimate and observations. *Mon. Wea. Rev.*, **132**, 1238-1253.

List of Tables

Table 1: The 10-year-avearged root-mean square analysis error and the associated default or tuned parameter values used in different data assimilation experiments for an assimilation window of 60 h (N=10) where R is the covariance localization radius, α is the covariance relaxation coefficient as in eq.9 and β is the mixing coefficient as in eq. 7. "NA" stands for not applicable and "failed" means no converged final analysis by that particular scheme.

		Ensemble size m = 40, default parameter setup		***Ensemble size m = 40, tuned parameter setup***		***Ensemble size m = 10, default parameter setup***		***Ensemble size m = 10, tuned parameter setup***	
		analysis error	**default R, α, β**	**analysis error**	**tuned R, α, β**	**analysis error**	**default R, α, β**	**analysis error**	**tuned R, α, β**
Perfect model F = 8.0	**4DVAR**	0.19	NA	0.19	NA	0.19	NA	0.19	NA
	EnKF	0.14	8, 0.5, NA	0.12	12, 0.3, NA	Failed	8, 0.5, NA	0.84	4, 0.7, NA
	E4DVAR1	0.13	8, 0.5, 1.0	0.11	12, 0.3, 1.0	0.13	8, 0.5, 1.0	0.13	8, 0.5, 1.0
	E4DVAR2	0.17	8, 0.5, 0.5	0.11	12, 0.3, 1.0	0.16	8, 0.5, 0.5	0.13	8, 0.5, 1.0
Moderate model error F = 8.5	**4DVAR**	0.45	NA	0.45	NA	0.45	NA	0.45	NA
	EnKF	0.68	4, 0.6, NA	0.64	3, 0.6, NA	Failed	4, 0.6, NA	1.48	3, 0.7, NA
	E4DVAR1	0.40	4, 0.6, 1.0	0.38	3, 0.6, 1.0	0.45	4, 0.6, 1.0	0.38	4, 0.7, 1.0
	E4DVAR2	0.36	4, 0.6, 0.5	0.35	3, 0.6, 0.4	0.40	4, 0.6, 0.5	0.36	4, 0.7, 0.3
Severe model error F = 9.0	**4DVAR**	1.12	NA	1.12	NA	1.12	NA	1.12	NA
	EnKF	Failed	4, 0.6, NA	1.24	3, 0.6, NA	Failed	4, 0.6, NA	1.76	2, 0.6, NA
	E4DVAR1	0.81	4, 0.6, 1.0	0.70	3, 0.6, 1.0	1.10	4, 0.6, 1.0	0.70	3, 0.7, 1.0
	E4DVAR2	0.80	4, 0.6, 0.5	0.66	3, 0.6, 0.4	0.88	4, 0.6, 0.5	0.68	4, 0.7, 0.3

Table 2: As in Table 1 but for an assimilation window of 24 h (N=4)

		Ensemble size m = 40, default parameter setup		*Ensemble size m = 40, tuned parameter setup*		*Ensemble size m = 10, default parameter setup*		*Ensemble size m = 10, tuned parameter setup*	
		analysis error	**default R, α, β**	**analysis error**	**tuned R, α, β**	**analysis error**	**default R, α, β**	**analysis error**	**tuned R, α, β**
Perfect model F = 8.0	**4DVAR**	0.39	NA	0.39	NA	0.39	NA	0.39	NA
	EnKF	0.14	8, 0.5, NA	0.12	12, 0.3, N	Failed	8, 0.5, NA	0.84	4, 0.7, NA
	E4DVAR1	0.14	8, 0.5, 1.0	0.12	12, 0.3, 0.8	0.14	8, 0.5, 1.0	0.14	8, 0.5, 1.0
	E4DVAR2	0.18	8, 0.5, 0.5	0.15	12, 0.3, 0.8	0.18	8, 0.5, 0.5	0.16	8, 0.5, 0.8
Moderate model error F = 8.5	**4DVAR**	0.77	NA	0.77	NA	0.77	NA	0.77	NA
	EnKF	0.68	4, 0.6, NA	0.64	3, 0.6, NA	Failed	4, 0.6, NA	1.48	3, 0.7, NA
	E4DVAR1	0.46	4, 0.6, 1.0	0.46	4, 0.6, 1.0	0.60	4, 0.6, 1.0	0.52	3, 0.5, 1.0
	E4DVAR2	0.42	4, 0.6, 0.5	0.41	3, 0.5, 0.4	0.44	4, 0.6, 0.5	0.42	4, 0.6, 0.3
Severe model error F = 9.0	**4DVAR**	1.52	NA	1.52	NA	1.52	NA	1.52	NA
	EnKF	Failed	4, 0.6, NA	1.23	3, 0.6, NA	Failed	4, 0.6, NA	1.74	2, 0.6, NA
	E4DVAR1	1.00	4, 0.6, 1.0	1.00	4, 0.6, 1.0	1.41	4, 0.6, 1.0	1.39	4, 0.7, 1.0
	E4DVAR2	0.86	4, 0.6, 0.5	0.86	4, 0.6, 0.5	1.09	4, 0.6, 0.5	1.01	4, 0.6, 0.3

List of Figures

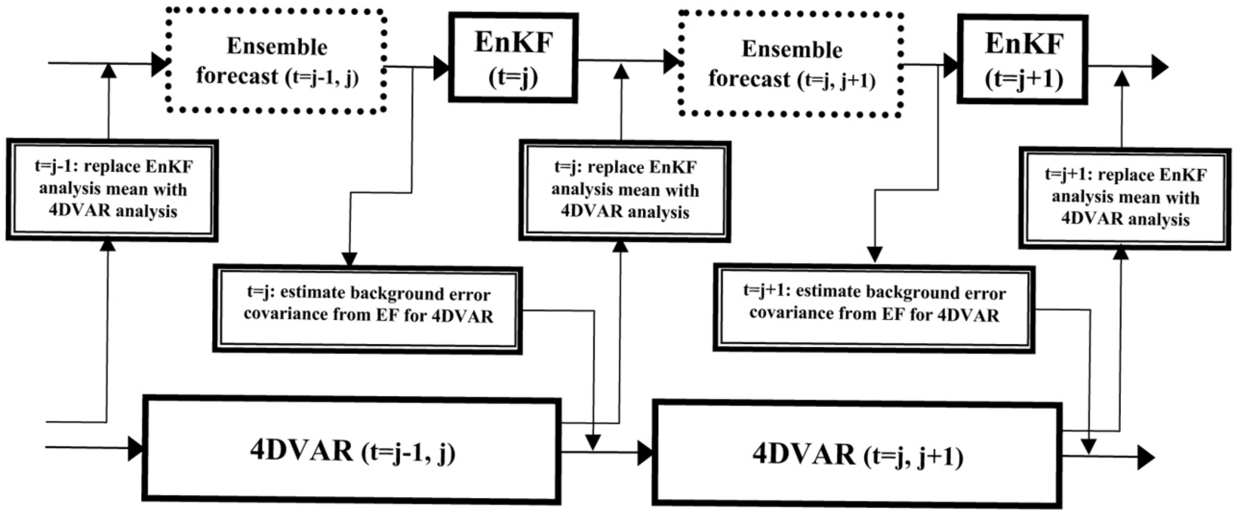

Figure 1: Schematics of the coupling between EnKF and 4DVAR that constitutes the E4DVAR used in this work.

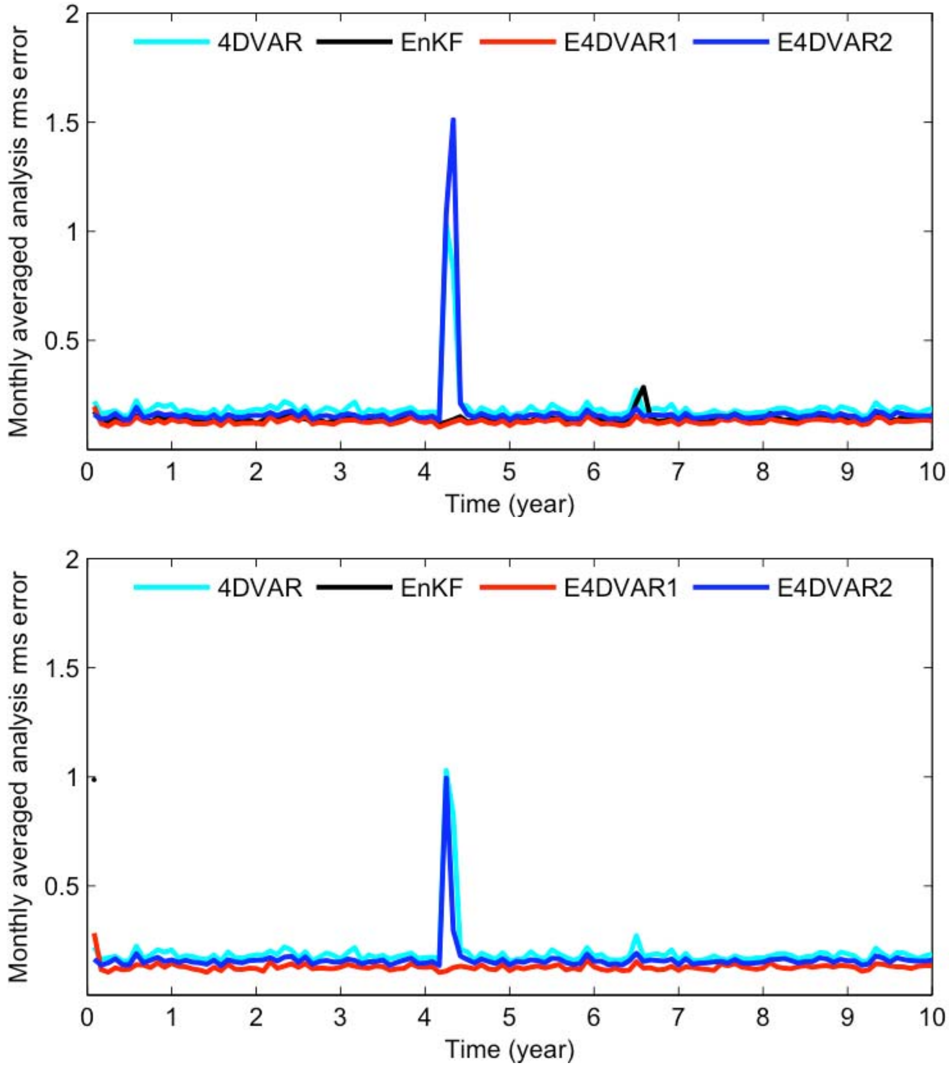

Figure 2: Time evolution of the monthly averaged root-mean square (rms) error for different data assimilation experiments with default parameter setups listed in table 1 for an assimilation window of 60 h (N=10) and am ensemble size m=40 (top) and m=10 (bottom) with a perfect forecast model (F=8.0). Some experiments may fail to converge to a solution and thus will not be plotted.

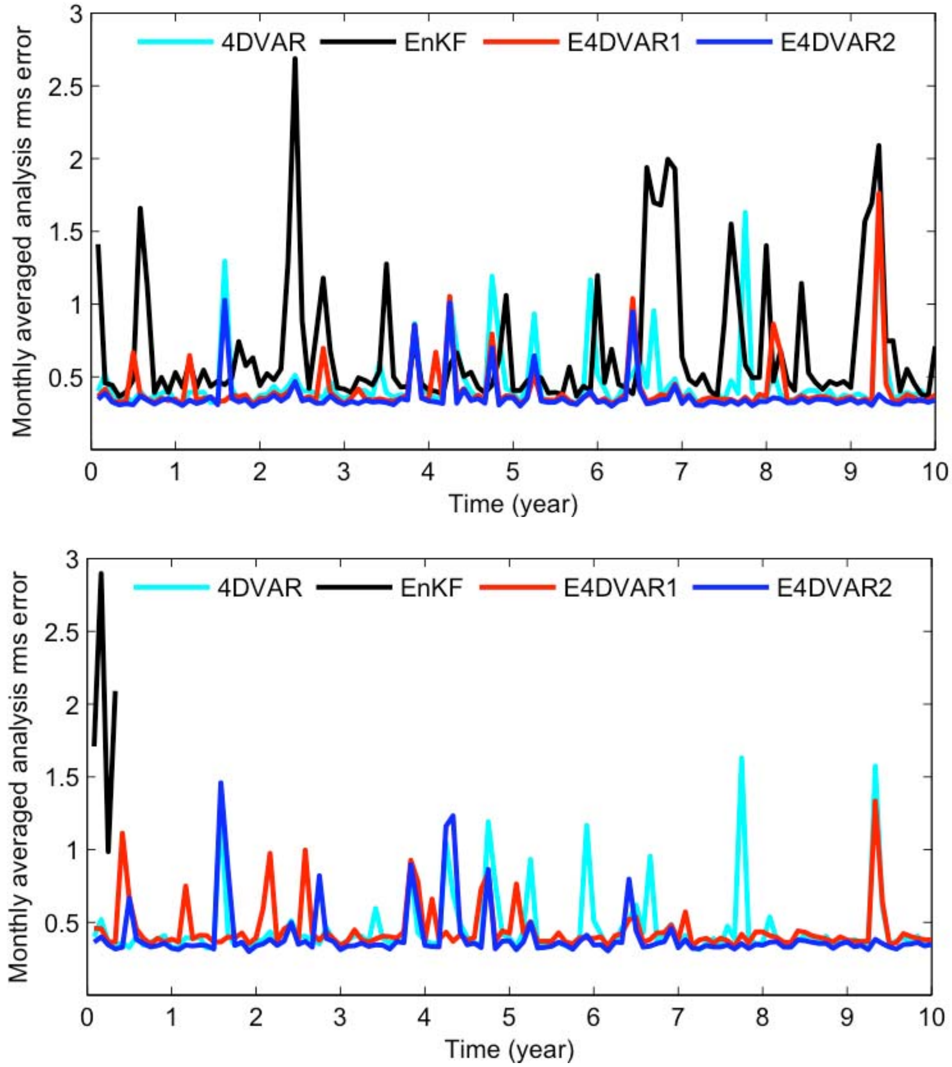

Figure 3: As in Figure 2 except for with moderate model error (F=8.5).

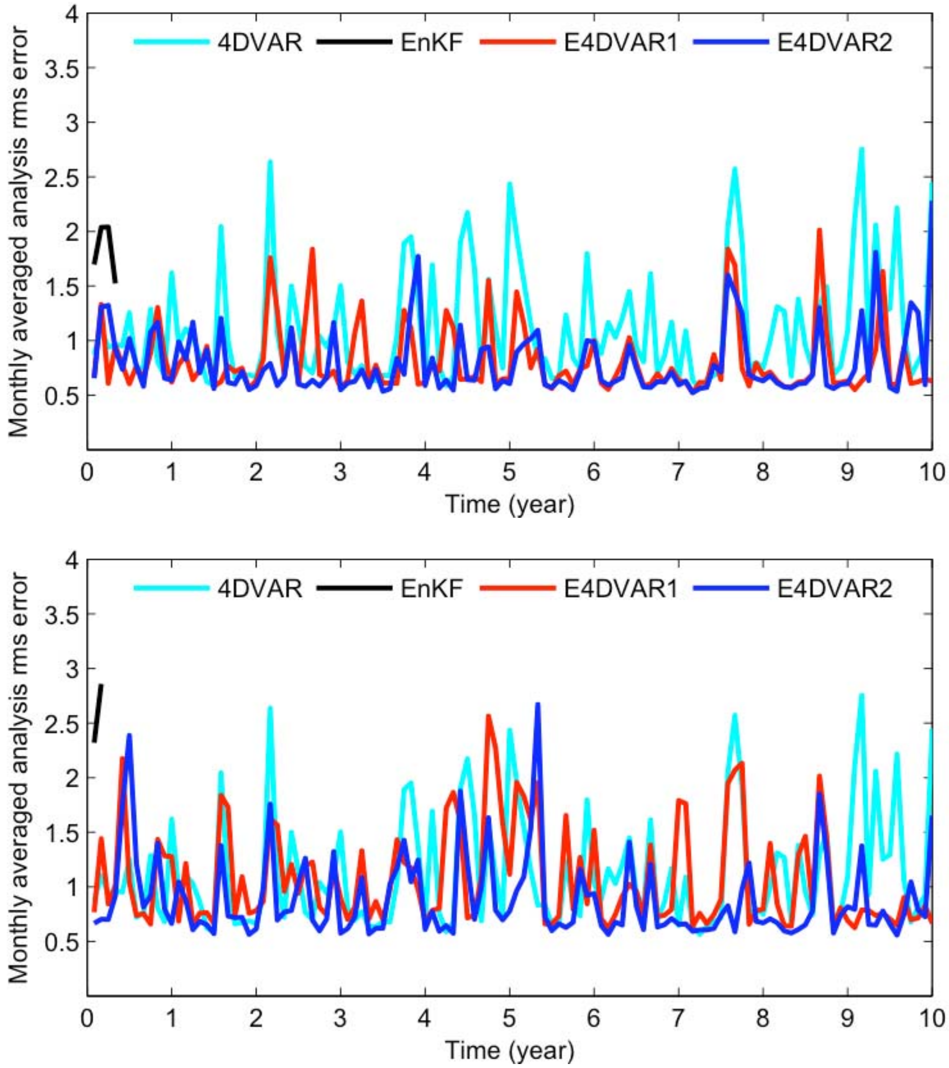

Figure 4: As in Figure 2 except for with severe model error (F=9.0).